Health IT infrastructure needs to support population health improvements in Colorado

Identity Management

Arthur Davidson, MD, MSPH

Denver Public Health

eHealth Commission

Office of eHealth Innovation (OeHI)
Colorado's State Designated Entity
Wednesday, May 11, 2016
303 17th Ave, Conference room 10A
Denver, CO

1

- Establish the problem(s)
- Review purpose and functions of a statewide
  - Master Patient Index (sMPI)
  - Master Provider Index (sMPrI)
- Present some next step options

Identity Management: No sure method to <u>know and uniquely identity</u> a client/patient/provider.

Client/Patient

- DMV
- Vital statistics
- CORHIO
- QHN
- APCD
- CIIS
- State OIT (justice, education, social services)

Provider

- DMV
- Institution A ….*n*
- Payer B ….*n*
- RCCO C…*n*
- DORA
- Medicare
- NPI
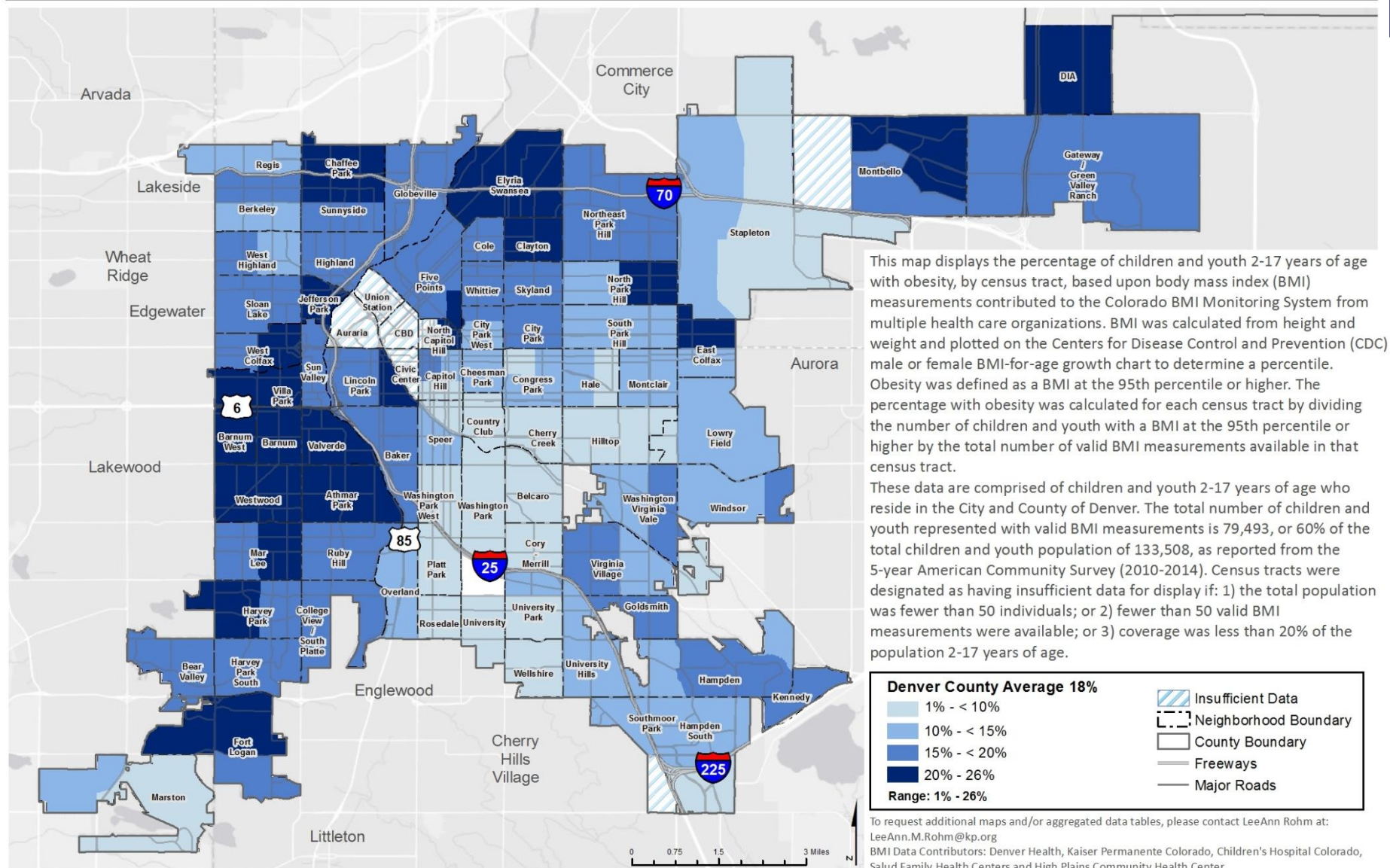
3

Statewide identity management is a 'team' sport…

SIM Advisory Board Orientation Packet (12/22/15)

- **Payment Reform:** Develop and implement **value based payment models** that incent integration and improve quality of care.

- **Practice Transformation**: Support practices as they accept new payment models and **integrate behavioral and physical** health care

- **Population Health**: Engage communities to reduce stigma, **promote prevention, and remove barriers** to accessing care
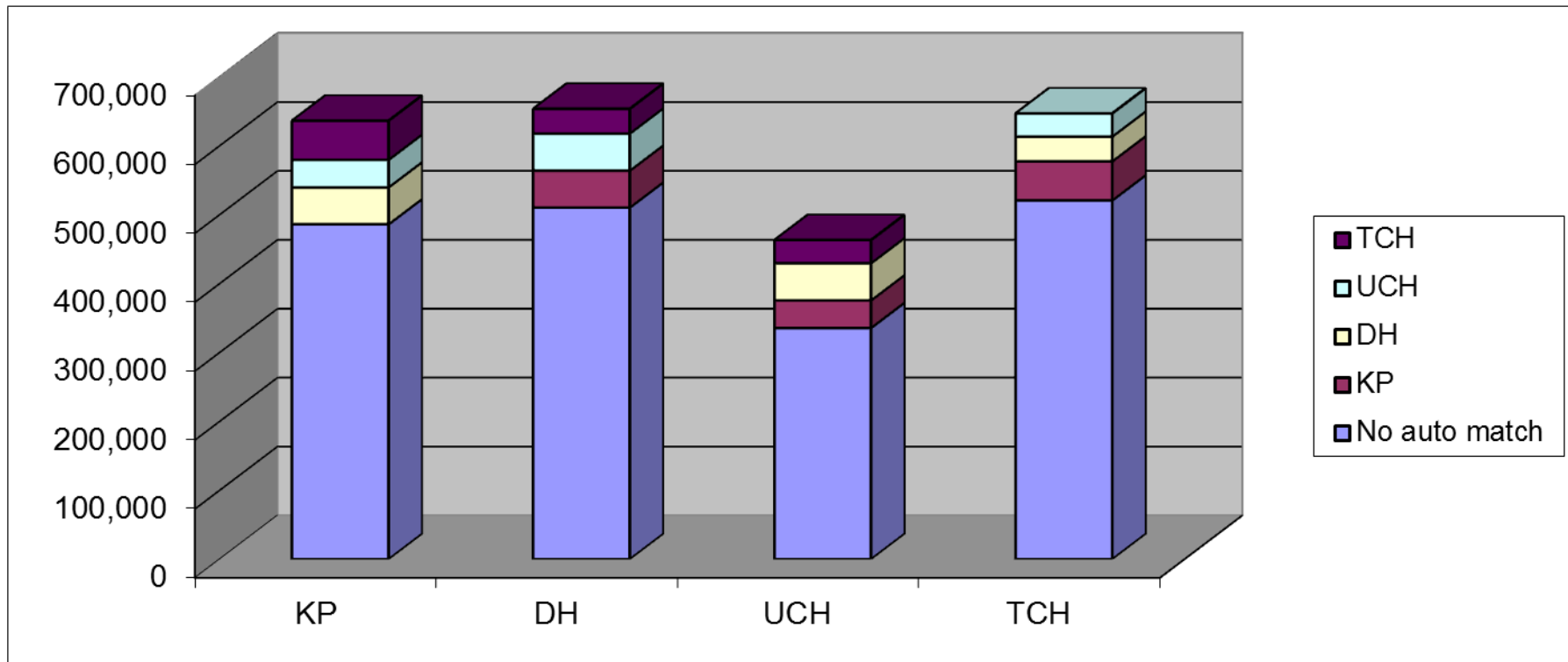
No team means ineffective identity management causing:

- **Failed attribution**

- **Inaccurate integration assessments**

- **Inaccurate monitoring**

4

Estimated Prevalence of Children and Youth with Obesity (2012-2014): City and County of Denver

This map displays the percentage of children and youth 2-17 years of age with obesity, by census tract, based upon body mass index (BMI) measurements contributed to the Colorado BMI Monitoring System from multiple health care organizations. BMI was calculated from height and weight and plotted on the Centers for Disease Control and Prevention (CDC) male or female BMI-for-age growth chart to determine a percentile. Obesity was defined as a BMI at the 95th percentile or higher. The percentage with obesity was calculated for each census tract by dividing the number of children and youth with a BMI at the 95th percentile or higher by the total number of valid BMI measurements available in that census tract.

These data are comprised of children and youth 2-17 years of age who reside in the City and County of Denver. The total number of children and youth represented with valid BMI measurements is 79,493, or 60% of the total children and youth population of 133,508, as reported from the 5-year American Community Survey (2010-2014). Census tracts were designated as having insufficient data for display if: 1) the total population was fewer than 50 individuals; or 2) fewer than 50 valid BMI measurements were available; or 3) coverage was less than 20% of the population 2-17 years of age.

Denver County Average 18%

- 1% - < 10%
- 10% - < 15%
- 15% - < 20%
- 20% - 26%

Range: 1% - 26%

- Insufficient Data
- Neighborhood Boundary
- County Boundary
- Freeways
- Major Roads

To request additional maps and/or aggregated data tables, please contact LeeAnn Rohm at: LeeAnn.M.Rohm@kp.org
BMI Data Contributors: Denver Health, Kaiser Permanente Colorado, Children's Hospital Colorado, Salud Family Health Centers and High Plains Community Health Center

Matching Across Institutions - *2009*

2,471,441 → **1,852,396 (75%)**

Problem – *Operational Plan (12/1/15)*

"The hub will…. *leverage the existing Master Patient Index (MPI), provider directories and other tools*. Building on clinical information, the phased approach will link to administrative claims information via the APCD and other sources as needed, providing a central aggregated clinical and cost data hub."

Identity Management - Functions

- Regular *automated receipt* of patient/client/provider identifying information from multiple partners
- Data are *standardized* for storage in the statewide master patient/provider index (sMPI/sMPrI)
- *Quality assurance* is performed on data with feedback to the data contributors/partners (e.g., remove duplicates)
- Process to *disambiguate records* is carried out (e.g., resolve potential overlaps across institutions)
- *Tools* are available for managing these processes and feedback to/from the data contributors/partner organizations

8

Disambiguation

A process of establishing a single semantic or meaning

- Matching process
- Resolves multiple potential matches
- Uses attributes of individuals (patients or providers) registered at multiple organizations

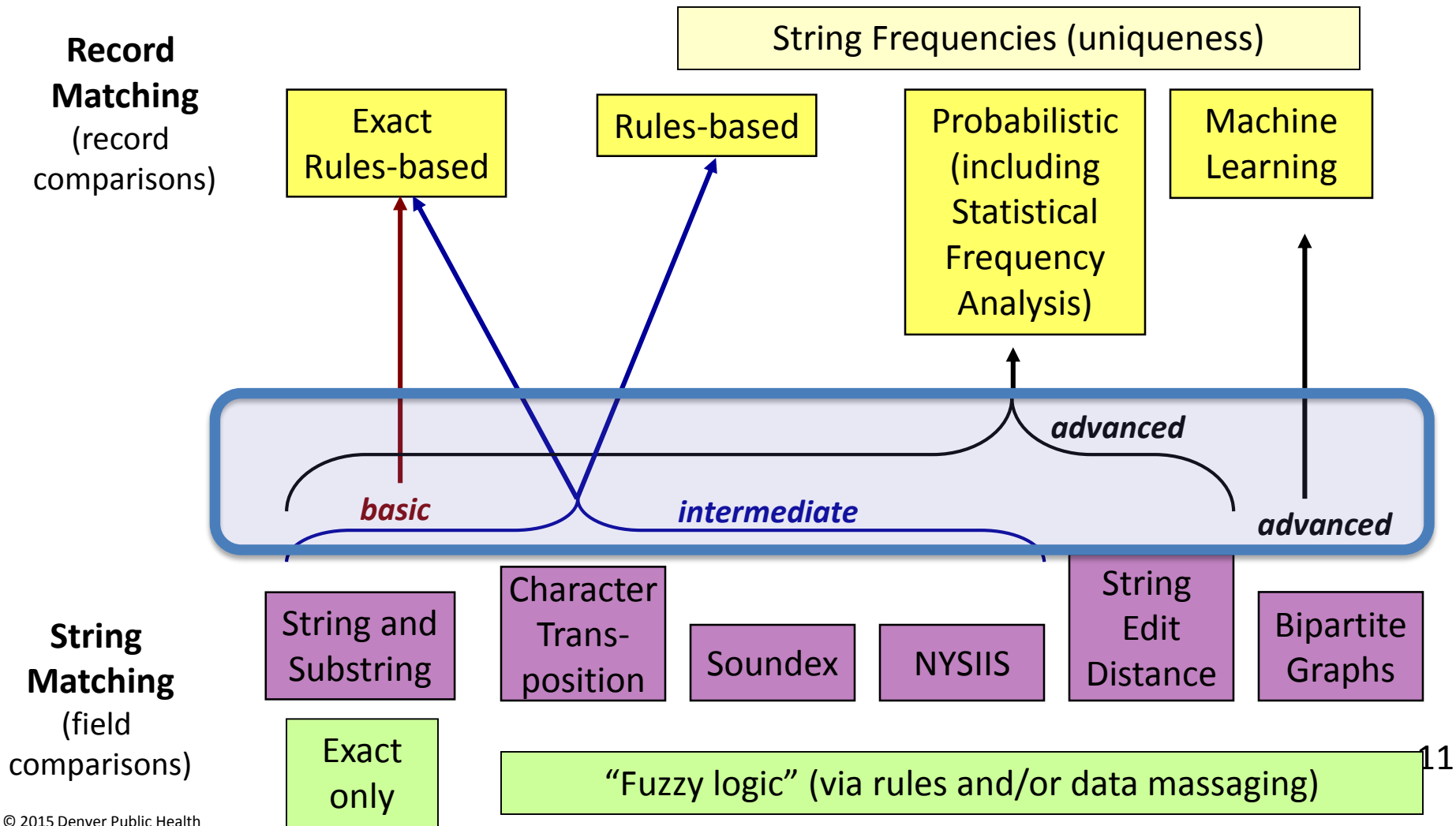***GOAL: find all matches for one target individual***

9

Deterministic vs. Probabilistic
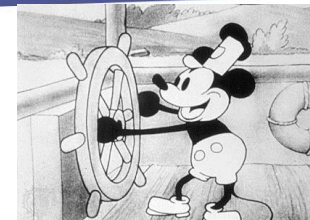
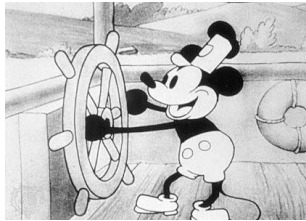**Deterministic indexing:** *perfect but inflexible matching*

- *False positives: none   False negatives: high*
- search based on an exact match of some combined factors (e.g., name, social security number, date of birth, and/or sex).
    - *Mickey Mouse, 11/18/28, M = Mickey Mouse, 11/18/28, M*

Probabilistic: *improves match by anticipating data entry errors/variance*

- *False positives: adjustable   False negatives: adjustable*
- rules-based search mechanism with some subset of exact matching
    - *Mickey Mouse, 11/18/28, M = **Mick** Mouse, 11/18/28, M*
    - *Mickey Mouse, 11/18/28, M = Mickey Mouse, 11/18/2**9**, M*
    - *Mickey Mouse, 11/18/28, M = **Micky** Mouse, 1**2**/18/28, M*

Field Transpositions

Data Range Comparisons

String Frequencies (uniqueness)

Record Matching
(record comparisons)

Exact Rules-based

Rules-based

Probabilistic (including Statistical Frequency Analysis)

Machine Learning

advanced

basic

intermediate

advanced

String Matching
(field comparisons)

String and Substring

Character Trans-position

Soundex

NYSIIS

String Edit Distance

Bipartite Graphs

Exact only

"Fuzzy logic" (via rules and/or data massaging)

© 2015 Denver Public Health

Errors in Linking

Electronic Linking Cause:

Mickey Mouse
DOB: 11/18/28

Mickey Mouse
DOB: 11/18/28

Records *seem* to match

Resulting error: **false positive** (overlay)
2 records linked under 1 MRN

Minnie Mouse
DOB: 05/15/28

Min**erva** Mouse
DOB: 05/15/**82**

Records *should* match

Resulting error: **false negative** (duplicate)
2 MRNs created

12

Compare the various data sources:

- a step-by-step procedure for solving a mathematical problem that frequently involves repetition of an operation especially using a computer

    - mathematical formula using a combination of weighted MPI/MPrI data elements to determine the probability of a **duplicate** or **overlap**

Same institution   *Different institutions*   13

Definitions

Duplicate Entry/File: *(undesirable and propagated)*

- more than one entry/file for the same patient or person (Rates around 9-15% ▪; 7-40% •)
  - *Mickey Mouse incorrectly has both record numbers 001 and 100 at Disneyland Clinic*
- may represent information capture errors

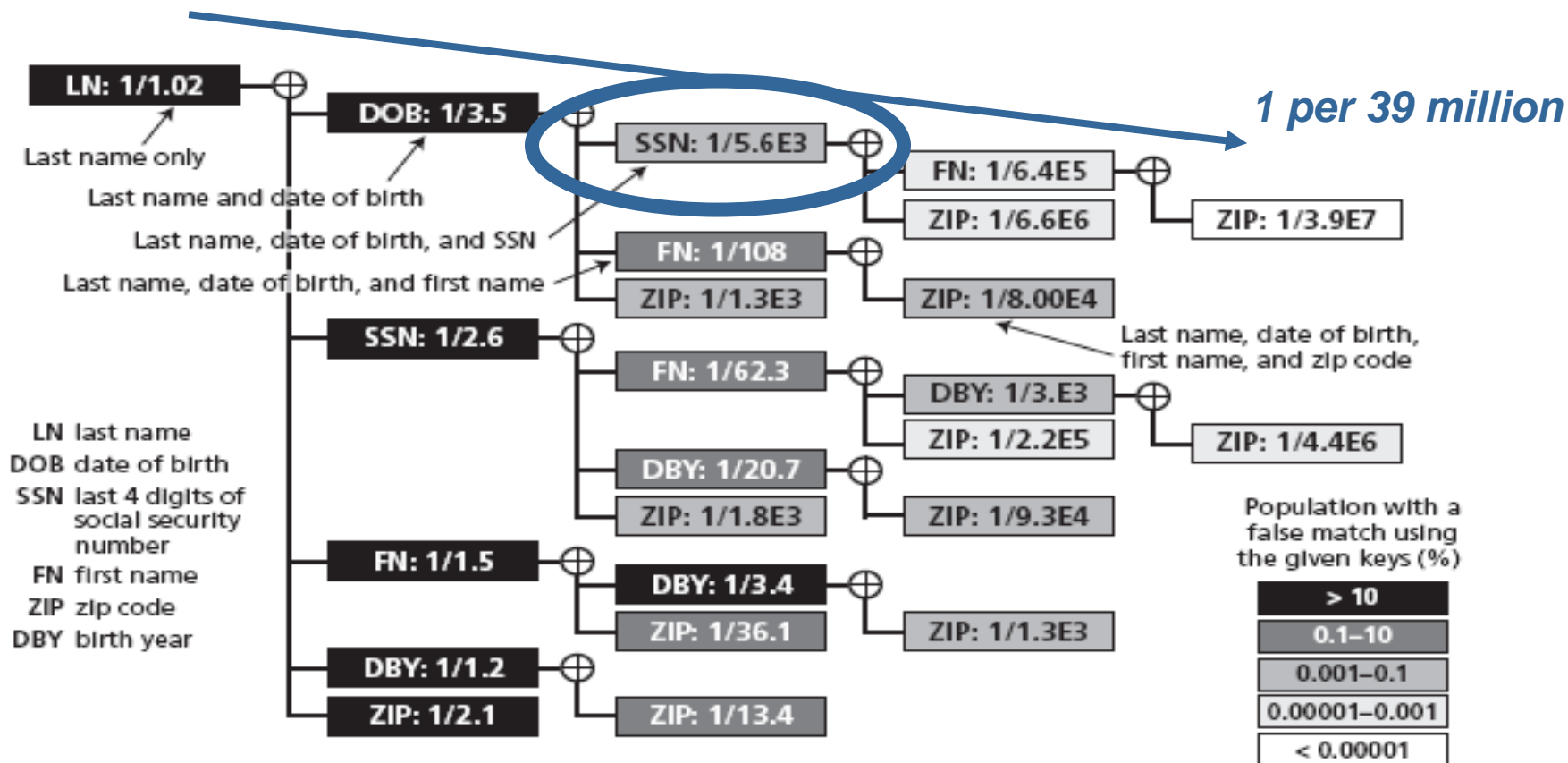**Overlay Entry/File:** *(undesirable and propagated)*

- more than one distinct individual assigned to the same record or identification number in a facility's MPI. (Among 2 hospital [n=5000] samples: 1 or 2 = rate of 0.02 – 0.04%)▪
  - *Mickey Mouse and Donald Duck incorrectly share record 001 at Disneyland Clinic*

Overlap Entries/Files: *(function of sMPI/sMPrI)*

- more than one MPI entry/file for the same patient/provider in two or more facilities across the state
  - *At Disneyland Clinic, Mickey Mouse has record 001 and record 100 at Disneyworld Clinic*
- algorithm works to identify and resolve overlaps without creating overlays

▪ *Grannis, Overhage, and McDonald 2004*
• *Initiate Systems, Inc. 2008*

14

1 per 39 million

LN: 1/1.02
Last name only

DOB: 1/3.5
Last name and date of birth

SSN: 1/5.6E3
Last name, date of birth, and SSN

FN: 1/108
Last name, date of birth, and first name

FN: 1/6.4E5
ZIP: 1/6.6E6

ZIP: 1/3.9E7

ZIP: 1/1.3E3

ZIP: 1/8.00E4

SSN: 1/2.6

FN: 1/62.3

DBY: 1/3.E3
ZIP: 1/2.2E5

ZIP: 1/4.4E6

Last name, date of birth, first name, and zip code

DBY: 1/20.7
ZIP: 1/1.8E3

ZIP: 1/9.3E4

LN last name
DOB date of birth
SSN last 4 digits of
   social security
   number
FN first name
ZIP zip code
DBY birth year

FN: 1/1.5

DBY: 1/3.4
ZIP: 1/36.1

ZIP: 1/1.3E3

DBY: 1/1.2
ZIP: 1/2.1

ZIP: 1/13.4

Population with a
false match using
the given keys (%)

| > 10 |
| 0.1–10 |
| 0.001–0.1 |
| 0.00001–0.001 |
| < 0.00001 |

Source: Social Security Death Master File
Note: Numbers in the blocks such as 1/1.02 in the leftmost dark block means there is 1
chance in 1.02 tires of a false positive match in database when this key type (LN) is used.
Moving to the right in the diagram 'DOB 1/3.5' means 1 chance in 3.5 tries of false positive
match when both last name and date of birth are used.

15

RAND MG753.22

Shared Content - *Hypothetical*

Data Element	DMV	CORHIO	QHN	APCD	…..
First Name	X	X	X	X	
Middle Name	X	X	X	X	
Last Name	X	X	X	X	
Gender (M / F / U)	X	X	X	X	
Address 1	X	X	X	X	
City	X	X	X	X	
State	X	X	X	X	
Zip	X	X	X	X	
County					
Phone	X	X	X	X	
SSN (Last 4 digits only)	X	?	?	?	
Date of birth	X	X	X	X	
Unique ID	X	X	X	X	
Payer number				X+	
MRN		X+	X+		

- Need reporting mechanisms to maintain sMPI/sMPrI accuracy
  - Cost-effective use of staff time as adds to collective accuracy
  - Established procedures:
    - Data sources:
      - receive/respond to potential duplicate reports
      - identify known non-duplicates
        - identify family members (twins, multi-generation name sharing)
        - share "never match" flag
    - sMPI/sMPrI:
      - rapidly remediate incorrectly matched records
  - Resolve intra-partner reports before determining overlaps

17

"… emphasis on a **collaborative approach** that shares data and insights across boundaries to drive better, more efficient medical practice and patient care."

"… drive the process of discovery as a natural outgrowth of patient care… to **ensure innovation, quality, safety, and value** in health care "

Institute of Medicine. 2007: http://www.iom.edu/Reports/2007/The-Learning-Healthcare-System-Workshop-Summary.aspx.

- ## Support (90/10 match) for HIE Architecture

"The free flow of information is hampered when not all doctors, facilities or other practice areas are able to make a complete circuit. Adding long-term care providers, **behavioral health providers**, and substance abuse treatment providers, for example, to statewide health information exchange systems will enable seamless sharing of a patients' health information between doctors or other clinicians when it's needed."

Slavitt/DeSalvo, https://blog.cms.gov/2016/03/02/bridging-the-healthcare-digital-divide-improving-connectivity-among-medicaid-providers/

Provider Directories: with an emphasis on dynamic provider directories that allow for bidirectional connections to public health and that might be web-based, allowing for easy use by other Medicaid providers with low EHR adoption rates

Development of a Master Patient Index (should be cost allocated)

Conclusions

- **To be successful CO needs identity management solutions:**
  - for SIM (near-term)
  - for proper attribution with payment reform
  - to accurately measure interventions and population health
  - to be a vibrant learning health system (long-term)
- **Identity management is complex**
  - need to establish robust tools and procedures
- **CO has an opportunity (and has been encouraged by CMS) to use 90/10 funding to build out a statewide:**
  - master patient index
  - master provider index

Questions/Discussion

adavidson@dhha.org